Materials and Methods

Ethics Statement

This study was approved by the Review Committee for IUPAC and permitted under MOU <insert\_permit\_number>.

**Sampling & DNA Extraction**

Tissue samples were collected from captive male and female fish 600 days post hatch (dph) located at the Fish Conservation and Culture Laboratory in Byron, CA. Muscle, gill, internal organ and scale tissues were sampled from each sex, the tissue was immediately flash frozen and stored on dry ice for transportation. High molecular weight (HMW) genomic DNA (gDNA) was extracted at the UC Davis DNA Technologies Core (Davis, CA) using the protocol described in Wasko et al. (2003).7 The size range of extracted gDNA fragments were determined using a pulse field gel run at a low frequency as to not shear the gDNA, for 24 hours. Physical sampling and extractions were repeated until the mean distribution of extraction lengths was 50kb or greater and there was sufficient quantity of HMW DNA to carry out each of the three required sequencing technologies (linked-read, long read and hi-c) for each sex (Supplemental Figure 1).

**Library Prep & Sequencing**

*Linked reads*

Once sufficient extracted fragments were acquired, genomic DNA was adjusted to a concentration of 0.91 ng/µl. Whole genome sequencing libraries were prepared using 10X Genomics Chromium Genome Library & Gel Bead Kit v.2 (10X Genomics, cat. 120258), Chromium Genome Chip Kit v.2 (10X Genomics, cat. 120257), Chromium i7 Multiplex Kit (10X Genomics, cat. 120262), and Chromium controller according to manufacturer’s instructions. After library preparation 1.14 ng of template gDNA was loaded on a Chromium Genome Chip and sequenced to roughly 80x coverage on an Illumina NovaSeq6000 150bp PE lane (Illumina, San Diego, CA).

*Long reads*

A PacBio HiFi SMRTbell® Libraries (<https://www.pacb.com/>) were prepped following the SMRTbell Express Template Prep Kit 2.0 procedure. The UC Davis Sequencing Center used a Megaruptor to shear DNA to an average length of ~15kb. The input for the library prep was 5.6ug of DNA and the library was size selected to 11kb. Prepped DNA was run on a Sequel II machine for a total of 30 hours.

*Hi-C Chromatin Confirmation Capture*

Hi-C library prep and subsequent sequencing were carried out at Phase Genomics in Seattle, WA.

**Quality Control & Genome Size Estimation**

*Illumina based reads (linked reads & hi-c)*

We used the software program kat6 and fastqc <cite> to perform quality control on raw Illumina files. First, to see if there is evidence of contamination, we observed the number of distinct k-mers at different frequencies for k-mers of length 16, 21, 27 and 33.

*Long reads*

asdf

**Assembly**

*Initial assembly from long reads*

We performed an initial phased genome assembly using the PacBio long reads with the Improved Phased Assembly (IPA) HiFi Genome Assembly software <cite>. Duplicated haplotigs were purged and the assembly was polished within the default parameters of the software.

*Scaffolding with linked-reads*

The initial phased assembly was then scaffolded using the 10X genomics Illumina linked-reads and the software scaff10x <cite>.

*Scaffolding with hi-c data*

We then mapped and formatted our Hi-C reads onto the scaffolded assembly using bwa-mem <CITE>, samtools <CITE>, picard <CITE>, bedtools <CITE> and custom perl scripts following the arima mapping pipeline <CITE>. After mapping the hi-c reads we performed one round of scaffoldind with SALSA2 <CITE>.

*Chromosome-level genome assembly using linkage map*

In order to achieve a chromosome level assembly we used linkage map data from Lew et al <CITE> to anchor scaffolds onto chromosomes using the Chromonomer software.

**Independent Chomosome Number Validation**

*Karyotype*

Asdf